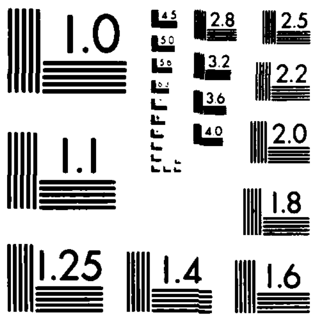
AD-A142 164 THE APPROXIMATE SOLUTION OF A SIMPLE CONSTRAINED SEARCH 1/
PATH MOVING TARGE..(U) NAVAL POSTGRADUATE SCHOOL
MONTEREY CA J N EAGLE MAY 84 NPS55-84-009
UNCLASSIFIED F/G 12/1 NL

END
DATE
FILMED
7-84
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

NPS55-84-009

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THE APPROXIMATE SOLUTION OF A SIMPLE CONSTRAINED SEARCH PATH MOVING TARGET PROBLEM USING MOVING HORIZON POLICIES

by

James N. Eagle

May 1984

Approved for public release, distribution unlimited.

Prepared for: Strategic Systems Project Office,
Arlington, VA 29376

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Commodore R. H. Shumaker
Superintendent

David A. Schrady
Provost

James N. Eagle
Associate Professor
Department of Operations Research

Reviewed by:

A. R. Washburn, Chairman
Department of Operations Research

Released by:

Kneale T. Marshall
Dean of Information and Policy
Sciences

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPS55-84-009	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) The Approximate Solution of a Simple Constrained Search Path Moving Target Problem Using Moving Horizon Policies		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) James N. Eagle		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93943		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS TASK 19422 N0003083WR32406
11. CONTROLLING OFFICE NAME AND ADDRESS Strategic Systems Project Office Arlington, VA 29376		12. REPORT DATE May 1984
		13. NUMBER OF PAGES 20
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

search
moving target

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A simple moving target problem is approximately solved for that search path minimizing the probability of non-detection subject to specified constraints on the search path. The solution procedure uses moving horizon policies. An "almost asymptotically geometric" structure is observed in the solution.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

S/N 0102- LF- 014- 6601

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

THE APPROXIMATE SOLUTION OF A SIMPLE CONSTRAINED SEARCH PATH MOVING TARGET PROBLEM USING MOVING HORIZON POLICIES

Presented here are the results of applying moving horizon policies to solve approximately a moving target problem, where both the searcher and the target have constraints on their paths. The solution procedure can be viewed as an approximation of the optimal dynamic programming method of Eagle (1982). This approximation may be useful if limits on available computer storage or computer time do not allow calculation of the optimal solution.

Only one problem geometry was examined. The problem was selected to keep the computer computations feasible rather than to be representative of any real-world search. It is possible that the patterns observed in the solution are specific to this problem geometry. Further work is required to establish the generality (or lack thereof) of these results.

1. The Problem

The target and searcher both move in discrete time among the 9 cells shown in Figure 1. The searcher starts in cell 1, and the target starts in cell 9. In each time period the searcher can move from his current cell to any adjacent cell. Cells are adjacent if they share a common side. The searcher can also choose to remain in his current cell. The target moves from cell to cell according to a specified Markov transition matrix. The probability of the target remaining in any cell i, given it was in cell i in the previous time period, is .4. The probability that the target transitions to any cell adjacent to i is $.6/c_i$, where c_i is the number of cells adjacent to i. So the target

transition matrix is

$$\begin{bmatrix} .4 & .3 & \emptyset & .3 & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset \\ .2 & .4 & .2 & \emptyset & .2 & \emptyset & \emptyset & \emptyset & \emptyset \\ \emptyset & .3 & .4 & \emptyset & \emptyset & .3 & \emptyset & \emptyset & \emptyset \\ .2 & \emptyset & \emptyset & .4 & .2 & \emptyset & .2 & \emptyset & \emptyset \\ \emptyset & .15 & \emptyset & .15 & .4 & .15 & \emptyset & .15 & \emptyset \\ \emptyset & \emptyset & .2 & \emptyset & .2 & .4 & \emptyset & \emptyset & .2 \\ \emptyset & \emptyset & \emptyset & .3 & \emptyset & \emptyset & .4 & .3 & \emptyset \\ \emptyset & \emptyset & \emptyset & \emptyset & .2 & \emptyset & .2 & .4 & .2 \\ \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & .3 & \emptyset & .3 & .4 \end{bmatrix}$$

If the searcher chooses the cell occupied by the target, then the target is detected with probability .5. If the searcher chooses a cell not occupied by the target, then the target can not be detected during that time period. The searcher has T time periods in which to search. His problem is to select that T-time period search path which minimizes the probability of target non-detection (PND).

1	2	3
4	5	6
7	8	9

Figure 1. 9-cell search grid.

2. Moving Horizon Policies

The problem presented was solved approximately using *m-time period moving horizon (m-TPMH) policies*. Such a policy is defined as follows: When T time periods remain in which to search and T > m, the m-TPMH policy selects as the next search cell that cell which would be optimal if m time periods remained in the problem. When T $\leq$ m, the optimal search path is selected. The 1-TPMH policy is called the *myopic policy*.

Moving horizon policies were introduced for the Markov decision process by Shapiro (1969) and have been recently suggested for search applications by Stewart (1984).

For this investigation, dynamic programming was used to construct the (m+1)-TPMH policy from the m-TPMH policy. The details are in Appendix A and Eagle (1982).

3. Experimental Results

A total of 320 cases were examined using problem lengths T (T=1,2,...,40) and m-TPMH policies (m=1,2,...,8). In addition, the optimal solutions were obtained (using dynamic programming and total enumerication) for T from 1 to 15 time periods. Figures 2 through 7 illustrate some observations suggested by the data collected.

Observation 1: For the moving horizon and optimal policies examined, the decrease in PND with increasing T was "almost asymptotically geometric."

Figures 2 through 5 illustrate "almost." In Figure 2, PND is plotted on a logarithmic scale against T. It appears here that PND for the myopic solution, the 8-TPMH solution, and the optimal solution are very nearly asymptotically geometrically decreasing. It is also apparent that the 8-TPMH policy generates a PND which decreases more rapidly than that generated by the myopic policy. Figures 3 and 4 show, however, that there is some fine structure in the graphs of PND which is not apparent in Figure 2. In Figure 3, the ratio PND(T)/PND(T-1) is plotted for the myopic and 8-TPMH policies. Figure 4 is a similar plot with an expanded y-axis scale. It appears that while the myopic policy is asymptotically geometric, the 8-TPMH policy is not. Graphs of PND(T)/PND(T-1) for the other moving horizon policies tested show an "almost asymptotically geometric" pattern similar to that of the 8-TPMH policy. (See Figures 5 and 6.)

Observation 2: It is possible for an m_1-TPMH policy to produce a smaller PND than a m_2-TPMH policy when $m_1 < m_2$.

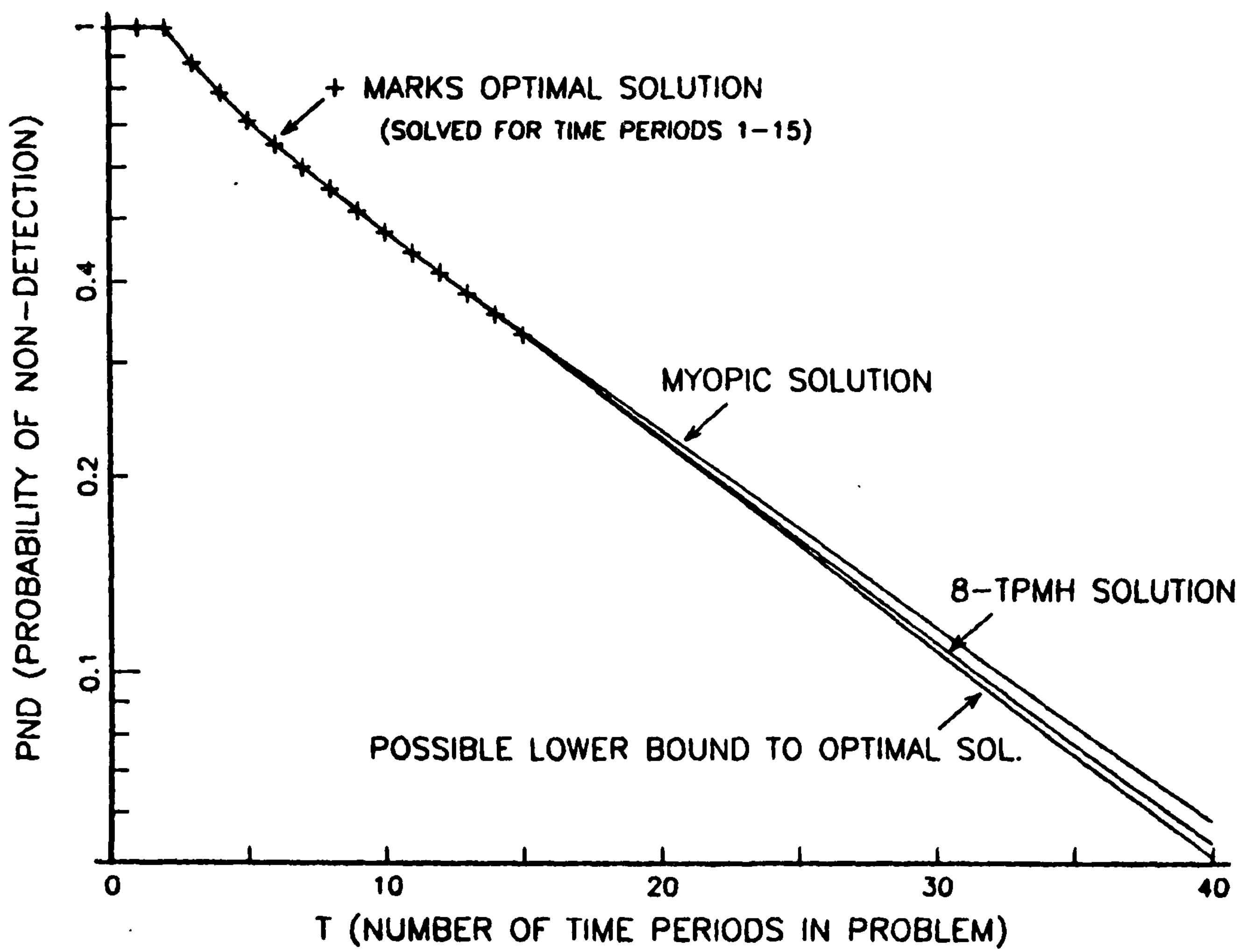


Figure 2. Probability of Non-Detection vs. Problem Length, T.

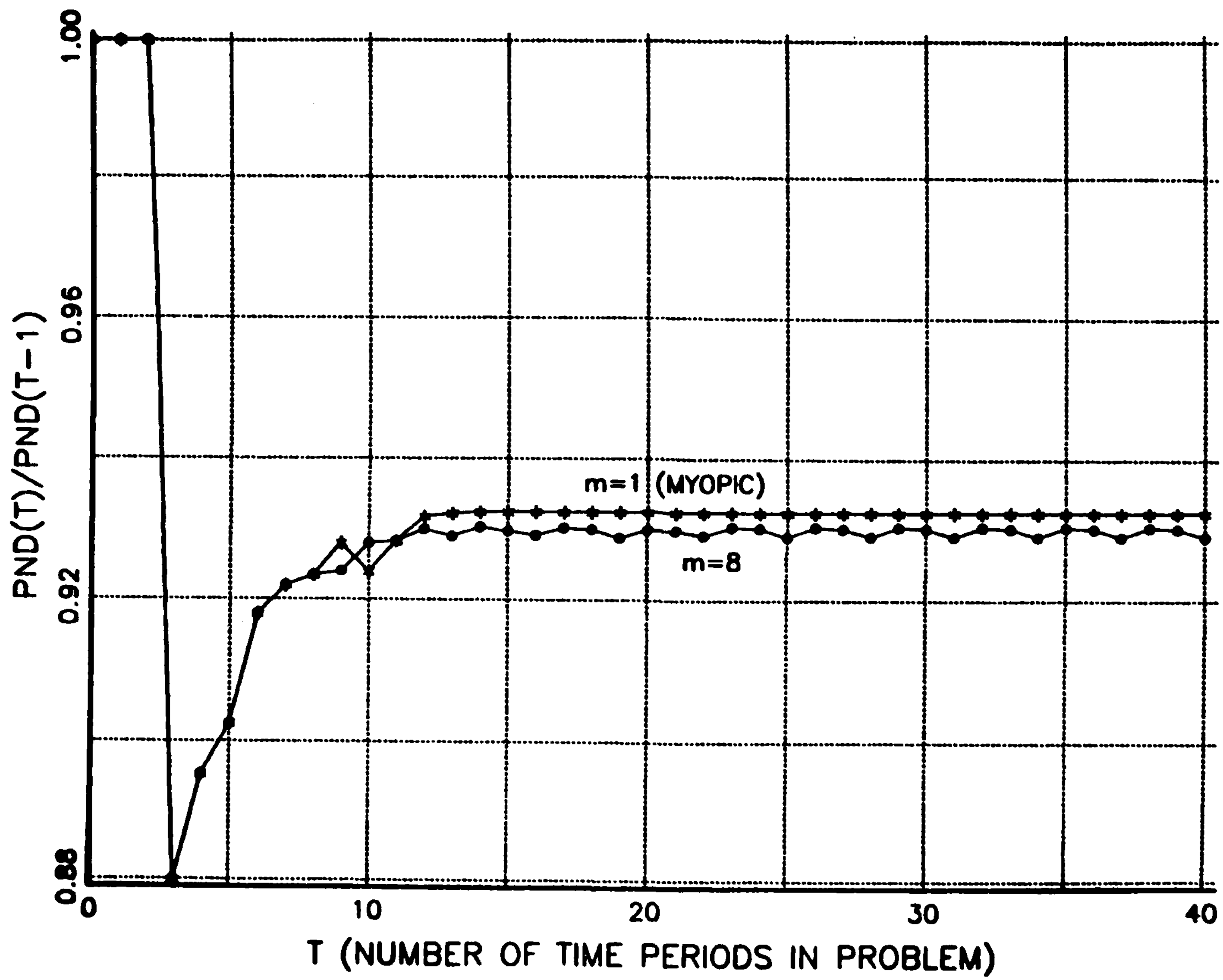


Figure 3. PDN(T)/PND(T-1) for Myopic and 8-TPMH Policies.

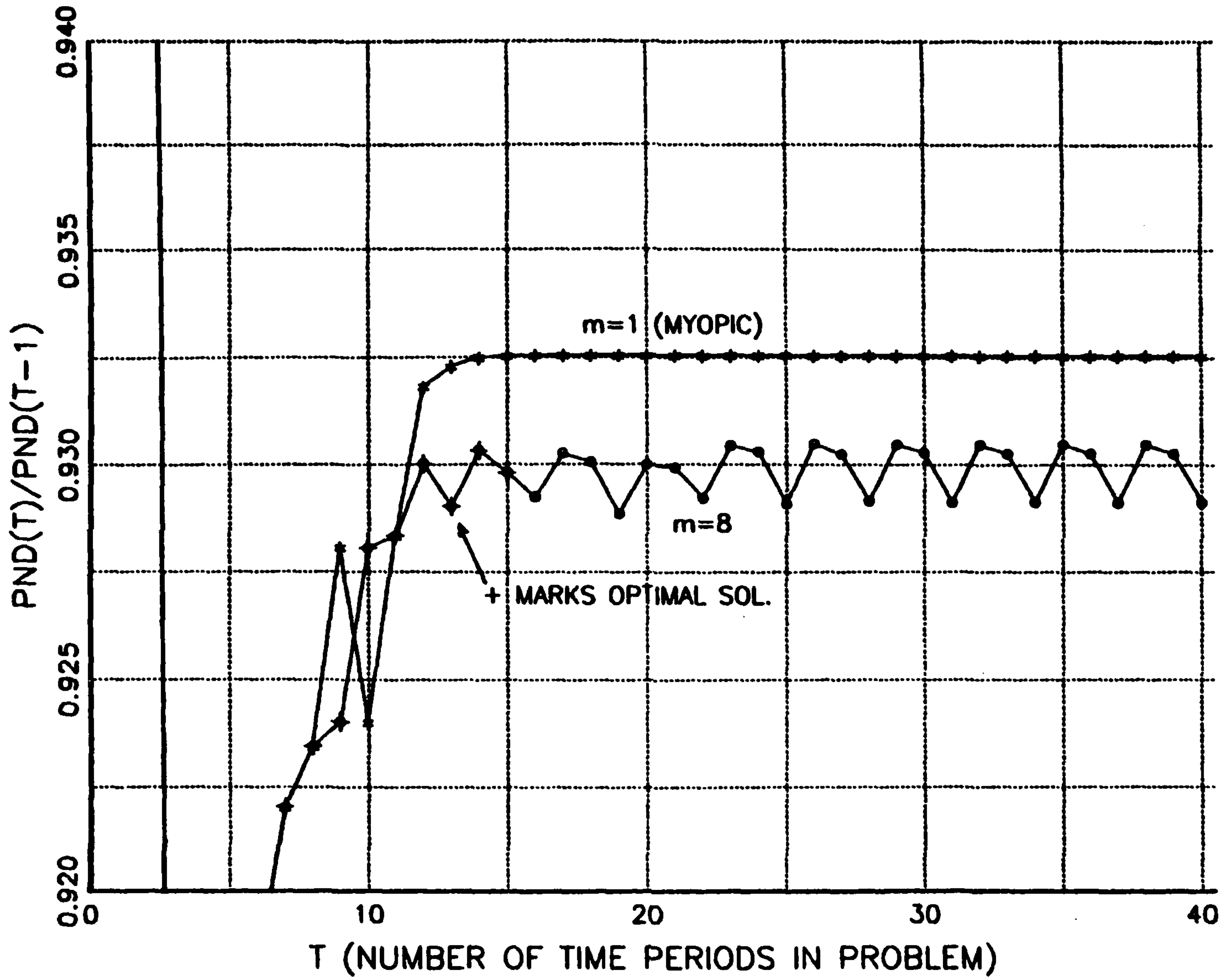


Figure 4. PND(T)/PND(T-1) for Myopic and 8-TPMH Policies.

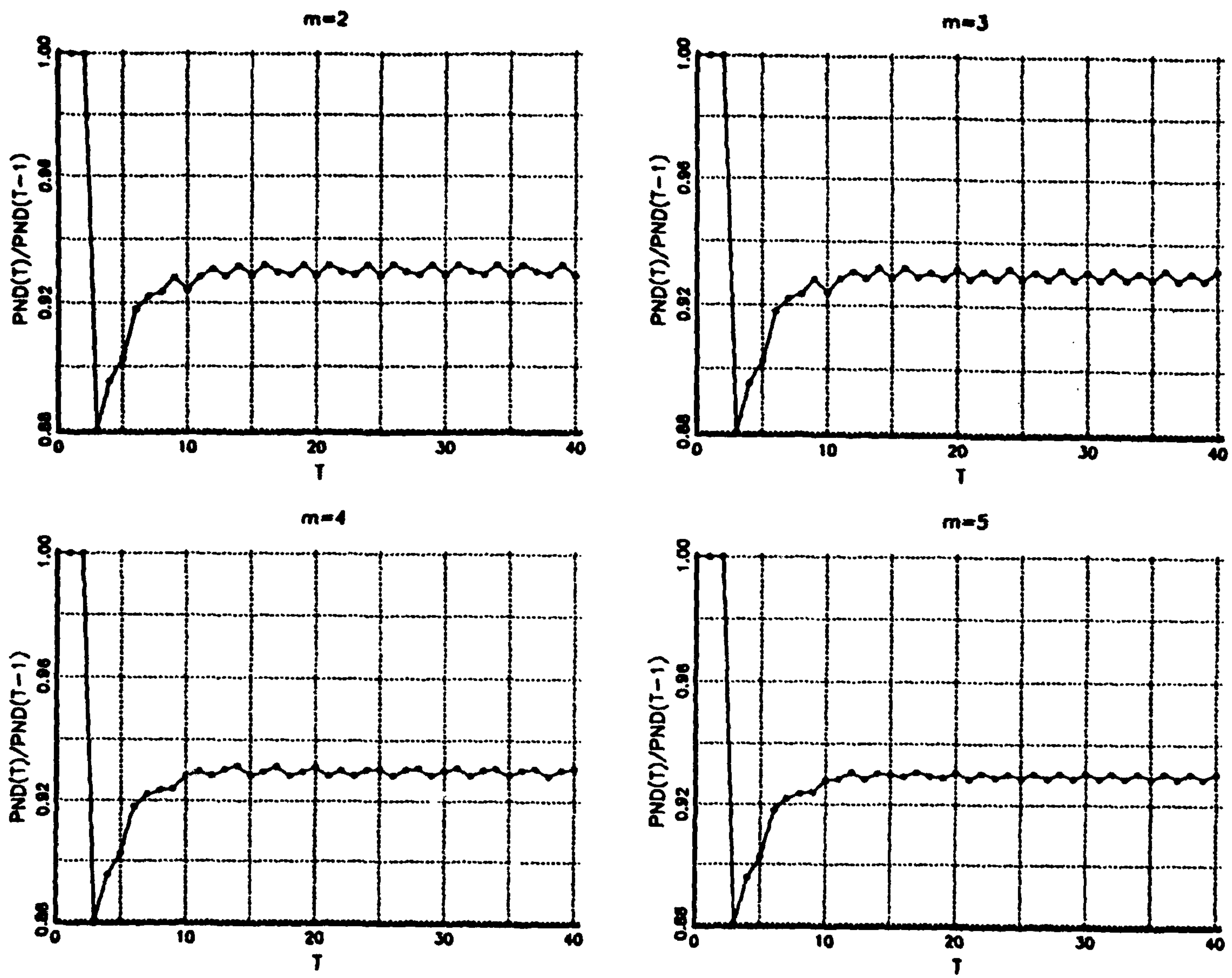


Figure 5. PND(T)/PND(T-1) for m-TPMH Policies (m = 2, 3, 4, 5).

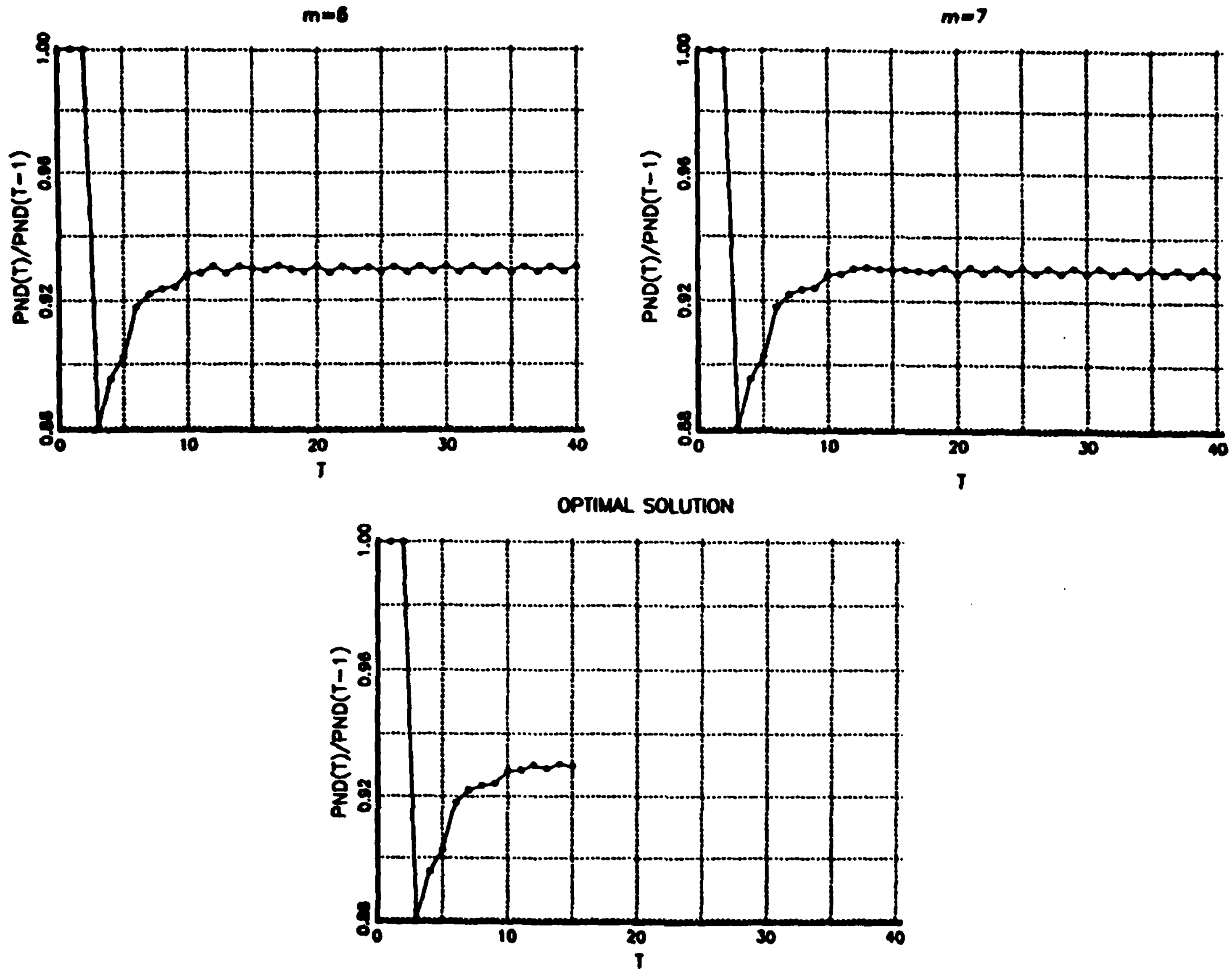


Figure 6. PND(T)/PND(T-1) for m-TPMH Policies (m = 6, 7) and Optimal Policy.

In general, m-TPMH policies performed better as m increased from 1 to 8, but there were some exceptions. Figure 7 illustrates. Here the difference in PND produced by the 3- and 4-TPMH policies is plotted against problem length T. A negative value of this difference indicates that the 3-TPMH policy performed better than the 4 TPMH policy for that particular value o. T. For example, for T=11, the 3-TPMH policy produced a PND of .4426, while the 4-TPMH policy gave .4434. The difference of -.0008 is plotted in Figure 7.

Observation 3: For $T \leq 15$, the optimal and 8-TPMH policies produced identical PND.

This is not to suggest that the 8-TPMH policy is optimal (It is not optimal. The 6-TPMH policy produced smaller values of PND for some T.), but rather that it may be a good approximately optimal policy for this problem.

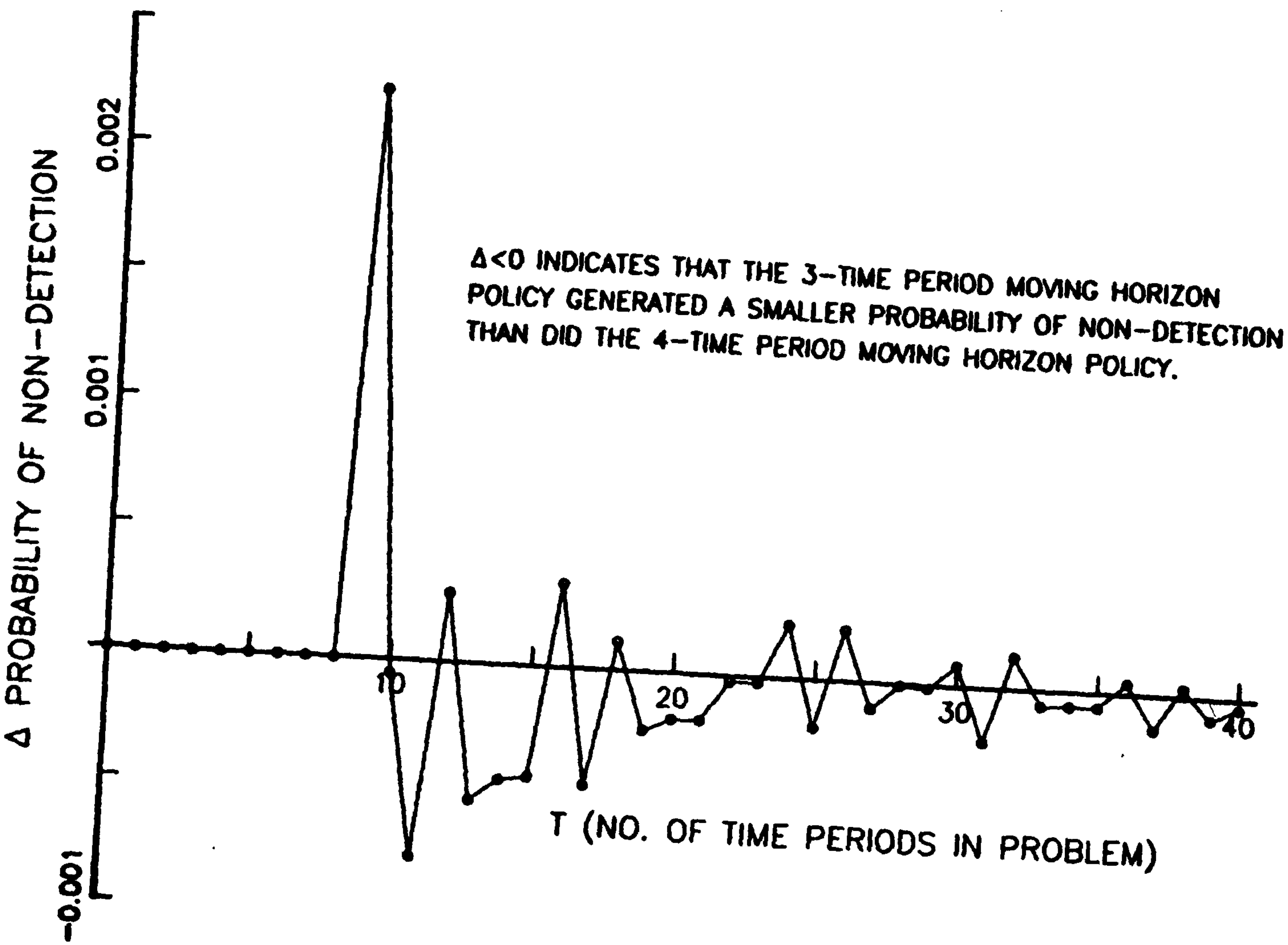


Figure 7. The Difference in PND Produced by 3-TPMH and 4-TPMH Policies.

4. Looking for a Lower Bound to PND

Moving horizon policies provide an upper bound to the optimal PND. It would be useful to construct a lower bound as well. If for all T greater than or equal to some $\hat{T}$, the optimal policy produced a non-decreasing PND(T)/PND(T-1) (as does the myopic policy in this example for $\hat{T} = 3$), then

$$PND(T) \geq PND(\hat{T}) \left(\frac{PND(\hat{T})}{PND(\hat{T}-1)} \right)^{(T-\hat{T})}$$

for all $T \geq \hat{T}$. Unfortunately, the optimal policy in this example did not generate non-decreasing PND(T)/PND(T-1). (See Figures 4 and 6.) The strongest statement about the optimal PND that the data collected can support is apparently the following:

For all $\hat{T} \in (1,2,\ldots,15)$ there exists a maximum $\gamma(\hat{T}) > 0$ satisfying

$$\hat{T} \leq T \leq 15 \Rightarrow \frac{PND(T)}{PND(T-1)} \geq \gamma(\hat{T}) \quad .$$

That is, for each $\hat{T}$, there was some maximum positive constant, $\gamma(\hat{T})$, which defined the tightest geometrically decreasing lower bound to PND(T), $T \geq \hat{T}$.

In addition, the data allow the following additional observation concerning the moving horizon PND.

<u>Observation 4</u>: For the m-TPMH policies examined with $T \geq 10$,

$$\frac{PND(T)}{PND(T-1)} \geq \frac{PND(10)}{PND(9)} \quad .$$

That is, for $T \geq 10$, the 1-time period geometric decrease in the moving horizon PND(T) was bounded below by PND(10)/PND(9). <u>If</u> this

observation also holds for the optimal policy, then for $T \geq 15$ we have for the optimal policy,

$$\begin{aligned} PND(T) &= PND(15)\ \frac{PND(16)}{PND(15)}\ \frac{PND(17)}{PND(16)}\ \cdots\ \frac{PND(T)}{PND(T-1)} \\ &\geq PND(15)\left(\frac{PND(10)}{PND(9)}\right)^{(T-15)} \\ &\geq .3308\ \ .9281^{\ (T-15)} \qquad (1) \end{aligned}$$

If (1) is a lower bound for this problem, it is a fairly tight one. This possible lower bound is plotted in Figure 2. Figure 8 shows the difference between this possible bound and the PND produced by the 8-TPMH, 2-TPMH and myopic policies. Figure 8 also suggests that increasing m from 1 to 2 resulted in considerably more policy improvement than did increasing m from 2 to 8.

5. Acknowledgement

Figures 2 through 8 were produced by an experimental APL package GRAFSTAT which the Naval Postgraduate School is using under a test agreement with IBM Watson Research Center, Yorktown Heights, N.Y. The author is grateful to Dr. P. D. Welch and Dr. Philip Heidelberger for making GRAFSTAT available.

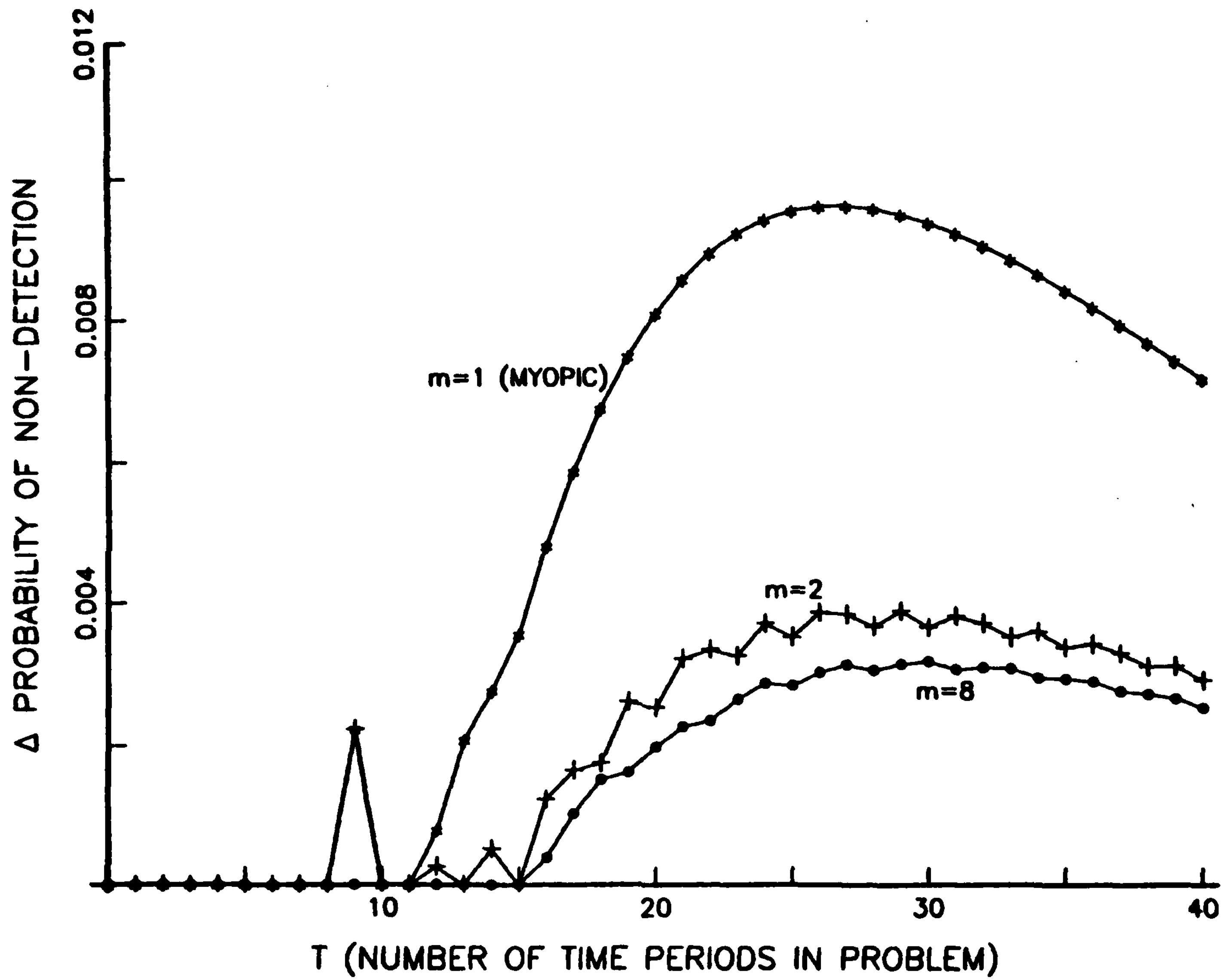


Figure 8. The Difference in PND Produced by m-TPMH Policies (m = 1, 2, 8) and a Possible Lower Bound to the Optimal Policy.

Appendix A: The Dynamic Programming Procedure for Determining Moving Horizon Policies

We make the following definitions:

C = set of all cells = $\{1,2,\ldots,N\}$,

C_j = set of all cells accessible in 1 time period to a searcher in cell j ,

q_j = P {target detection|target in cell j and search conducted in cell j} ,

p_{ij} = P {target transitions in 1 time period from cell i to cell j} ,

P = target transition matrix = $[p_{ij}] \in R^{N\times N}$,

d_n = the cell searched when n time periods remain in the problem,

$\delta^n = (d_n, d_{n-1},\ldots,d_1)$ = an n-time period search path,

π_j = probability that the target is in cell j ,

$\pi = (\pi_1,\pi_2,\ldots,\pi_N)$ = target probability distribution over C.

With any n-time period search path, δ^n, there can be associated a vector $a \in R^N$ such that a_i = P{target detection|δ^n is followed; target in cell i when search begins}. The probability of detection when δ^n is followed and the initial target distribution is π is then πa. Now let A(n,i) be the set of vectors associated with all possible δ^n, given the searcher is in cell i when n time periods remain. Then the maximum obtainable n-time period probability of detection given an initial target distribution of π is

$$V_n(\pi,i) = \max_{a \in A(n,i)} \pi a \quad . \qquad \text{(A1)}$$

And the optimal n-time period search path is that δ^n associated with the maximizing $a \in A(n,i)$.

The dynamic programming problem is then to construct the vector sets $A(n+1,1)$, $A(n+1,2),\ldots,A(n+1,N)$ from the vector sets $A(n,1)$, $A(n,2),\ldots,A(n,N)$. Also, each $a \in A(n+1,i)$ must have associated with it an $(n+1)$-time period search path.

Let $\hat{a}$ be any element of $A(n,j)$ and δ^n be the n-time period search path associated with $\hat{a}$. Now the N-vector associated with the $(n+1)$-time period search path (j,δ^n) is

$$a = e_j\, q_j + P_j\, \hat{a} \quad ,$$

where $e_j \in R^N$ is the j-unit vector and $P_j \in R^{NxN}$ is P with row j multiplied by $(1-q_j)$. To see this, the components of a and $\hat{a}$ are interpreted as probabilities of detection when n+1 and n searches respectively remain in the problem. The entire set $A(n+1,i)$ is then

$$\{a \in R^N | a = e_j\, q_j + P_j\, \hat{a}\ ;\ j \in C_i\ \&\ \hat{a} \in A(n,j)\}. \quad (A2)$$

The dynamic programming process begins by setting $A(0,i) = 0 \in R^N$, $i = 1,2,\ldots,N$. One iteration gives the myopic solution. Specifically, applying (A2) when $A(0,i) = 0$ yields

$$A(1,i) = e_i\, q_i,\ i = 1,\ldots,N \ ,$$

with an associated 1-time period search path of $\delta^1 = d_1 = i$. Continued application of (A2) allows recursive construction of the sets $A(n,i)$ with an n-time period search path associated with each vector in each set.

The set $A(n,i)$ constructed in this manner from the sets $A(n-1,j)$, $j \in C_i$, may contain some vectors which will never maximize (A1) for any target distribution π. The δ^n associated with each of these "dominated" vectors can not be an optimal n-time period search path. To test whether a vector $\hat{a} \in A(n,i)$ is dominated, the following linear program is solved:

$$\min_{\pi, x} \quad x - \pi\hat{a}$$

$$\text{s.t.} \quad x \geq \pi a, \; a \in A(\hat{a})$$
$$\pi \in \Pi$$

where $A(\hat{a})$ is the set $A(n,i)$ less the vector $\hat{a}$, and $\Pi = \{\pi \in R^N | \pi_i \geq 0 \text{ and } \sum_i \pi_i = 1\}$. Whenever the minimal value of $x - \pi\hat{a}$ is non-negative, $\hat{a}$ is dominated and can be removed from $A(n,i)$. Only the non-dominated vectors in $A(n,i)$ need be used to construct $A(n+1,j)$. Letting B be the convex hull of $A(\hat{a})$, Eagle (1982) showed that $\hat{a}$ is dominated if and only if there exists some $b \in B$ such that $b \geq \hat{a}$.

A simpler domination procedure is to remove $\hat{a}$ from $A(n,i)$ wherever there exists a vector $a \in A(\hat{a})$ such that $a \geq \hat{a}$. This method is easier to implement than the linear programming procedure, but does not reduce $A(n,i)$ to its minimum size. Thus more computer storage is required to save $A(n,i)$ in each stage of the dynamic program.

Once the vector sets $A(m,i)$, $i = 1,\ldots,N$, have been constructed and a δ^m has been associated with each $a \in A(m,i)$, then the m-TPMH policy is available. Assume $n > m$ time periods remain in the problem, the searcher is in cell i, and the target

distribution is π. Then the m-TPMH policy picks as d_n the first element of δ^m, where δ^m is the m-time period search path associated with

$$\underset{a \in A(m,i)}{\operatorname{argmax}} \, \pi a \, . \tag{A3}$$

If the target is not detected in time period n, the target distribution given a Bayesian update for the unsuccessful search and (A3) is used again to determine d_{n-1}. When the problem solution progresses to the point where m time periods remain, the m-TPMH policy picks the optimal δ^m for the remaining time periods.

REFERENCES

EAGLE, J. N., 1982. The Optimal Search for a Moving Target When the Search Path is Constrained. Naval Postgraduate School Technical Report NPS55-83-013.

SHAPIRO, J. F., 1969. Turnpike Planning Horizons for a Markovian Decision Model, Management Science, 14, pp. 1635-1660.

STEWART, T. J., 1984. Optimizing Search with Positive Information Feedback, National Research Institute for Mathematical Science Technical Report TWISK 313, Pretoria, South Africa.

INITIAL DISTRIBUTION LIST

	No. Copies
Library, Code 0142 Naval Postgraduate School Monterey, CA 93943	4
Director of Research Administration Code 012A Naval Postgraduate School Monterey, CA 93943	1
Library, Code 55 Naval Postgraduate School Monterey, CA 93943	1
Defense Technical Information Center Cameron Station Alexandria, VA 22314	1
J. N. Eagle, Code 55Er Naval Postgraduate School Monterey, CA 93943	20

TE

MED

8